H 81

NCHSR

National Medical Expenditure Survey

Sample Design of the National Medical Expenditure Survey - Household Component

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
Public Health Service
National Center for Health Services Research
and Health Care Technology Assessment

HE 20.6517:

READEX FICHE best copy available

SAMPLE DESIGN OF THE NATIONAL MEDICAL EXPENDITURE SURVEY-HOUSEHOLD COMPONENT

Steven B. Cohen

National Center for Health Services Research
and Health Care Technology Assessment
Parklawn Building, Room 18A-55
5600 Fishers Lane
Rockville, Maryland 20857

Ralph DiGaetano
Joseph Waksberg

Westat, Incorporated
1650 Research Boulevard
Rockville, Maryland 20854

Presented at the Annual Meetings of the American Statistical Association, San Francisco, California.

The views expressed in this paper are those of the authors and no official endorsement by the Department of Health and Human Services or the National Center for Health Services Research and Health Care Technology Assessment is intended or should be inferred. The author wishes to thank Ms. D.E.B. Potter, Dr. David Fitch and Mr. Sherman Edwards for their helpful comments and to thank Ms. Evelyn Manstof and Mr. Bill Ashby for their skillful typing of this manuscript.

SAMPLE DESIGN OF THE NATIONAL MEDICAL EXPENDITURE SURVEY-HOUSEHOLD COMPONENT

Steven B. Cohen, Ralph DiGaetano and Joseph Waksberg

Introduction

The Household Component of the National Medical Expenditure Survey (NMES) was established to provide an assessment of the health care utilization, costs, sources of payment and health insurance coverage of the U.S. civilian noninstitutional population. The period of assessment covers calendar year 1987, during which data are collected from a national probability sample of dwelling units. The survey was designed to provide data for a major research effort in the Division of Intramural Research of the National Center for Health Services Research and the Health Care Technology Assessment (NCHSR), and was cosponsored with the Health Care Finance Administration (HCFA). The data will meet the needs of government agencies, legislative bodies, and health professionals for more comprehensive national data required for the analysis and formulation of national health policies. NMES succeeds a series of national medical expenditure surveys, most notably the 1980 National Medical Care Utilization and Expenditure Survey (NMCUES) and the 1977 National Medical Care Expenditure Survey (NMCES). Unlike the previous efforts, the NMES household sample of approximately 16,000 dwelling units includes the oversampling of groups of particular policy interest: blacks, Hispanics, the poor and near poor, the elderly and persons with functional limitations.

The NMES household survey is a year long panel, collecting measures of health status, use of health care services, expenditures and sources of payment, insurance coverage, employment, income and assets, as well as demographic information. An initial screening interview was conducted in the fall of 1986 for a sample of approximately 35,000 addresses, to obtain information required for the oversampling of specific policy relevant population subgroups. Field operations for the NMES household component consist of four core interviews conducted with the selected households at three to four month intervals over a fifteen month period. The first two interviews will be conducted in person, the third by telephone (if a telephone is available and the procedure is acceptable to the household), and the final core interview will again be in person.

The sample design can be characterized as a stratified multi-stage area probability design. The survey was conducted from two independently drawn national samples selected from two distinct survey organizations, Westat, Inc., and the National Opinion Research Center (NORC). This paper provides a detailed description of the NMES household survey design (HHS), sample yields from the screener interview and a discussion of survey design complexities which require special consideration for analysis.

The NMES Sample Design

An examination of the analytical goals of the study and budget constraints indicated that the sample design for the NMES household survey should satisfy the following requirements:

1. The full series of interviews should be completed in approximately 14,000 households.

2. The sample should be spread over at least 100 separate areas to represent the civilian non-institutionalized population of the 50 states and the District of Columbia.

3. The sample shall be designed to produce unbiased national estimates and unbiased estimates with adequate precision for the four census regions.

4. The sample shall be designed to meet predesignated precision specifications for the following population subgroups of analytical interest: blacks, Hispanics, the poor and near poor, the elderly (65+), and the functionally impaired. It was recognized that in order to achieve the NMES precision requirements for the overall population and specified domains of interest, an initial sample selection substantially larger than 14,000 households with subsequent screening would be required.

The numerical constraint on a minimum number of geographical localities was specified to ensure sufficient geographic dispersion of the sample and allow for separate regional estimates. The precision specifications were provided to insure that the design would meet analytical objectives and to facilitate stage-specific sample size determination.

The adopted NMES household survey sample design is a stratified area probability design with four stages of sample selection: (1) selection of primary sampling units (PSU's), which are counties, parts of counties or groups of contiguous counties; (2) selection of segments within PSU's; (3) selection and screening of dwelling units within segments; and (4) selection of dwelling units based on demographic characteristics (both household and individual level) from the set of screened dwelling units.

The sample of PSU's represents a union of the national sample frames of Westat, Inc., and the National Opinion Research Center (NORC). The national general purpose area samples from Westat and NORC that comprise the NMES household sample are similar in structure, thus simplifying the development of sampling, listing and interview procedures and permitting a single management control system. Since the sampled PSU's, segments, and dwelling units selected for screening were representative of two independently drawn national samples, except for difficulties associated with survey nonresponse and other non-sampling errors, statistically unbiased national and domain specific estimates can be produced from each sample or from the two samples combined. The combined sample includes 165 PSU's located in 127 distinct sites. The number of separate primary areas is less than the total primary sampling units in the two national samples because some areas are in both samples.

More specifically, the Westat first stage sample was stratified by social, economic and demographic characteristics which included region, SMSA status, percent of population employed, percent white, and percent over age 65. In all, 81 PSU's were selected for NMES. Similarly, the NORC first stage sample included the following stratification measures: region, SMSA status, and population size. The NMES sample consisted of 84 NORC PSU's. Within PSU's, a two or three stage sample design was used to select dwelling units for the

screening sample. The first stage consisted of 1980 Census Enumeration Districts (ED's) or individual block or block combinations. The second stage was only used when ED's or block groups were exceptionally large in area or number of households, respectively. Then "chunking" or the partitioning of the ED or block group was employed, dividing the selected area into several smaller segments of approximately equal size in terms of households, one of which was randomly selected. The third stage consists of the selection of all or a systematically selected subsample of households within each area segment. The ED's or blocks were selected with probability proportionate to size, with a systematic procedure allowing for implicit geographic stratification. The sampling and subsampling rates were specified so that all dwelling units in the U.S. had an equal probability of selection.

Within the sampled PSU's, 2,317 segments were selected (1,150 for Westat, 1,167 for NORC). The segment sampling process resulted in a set of maps showing the boundaries of the sampled segments and their associated probabilities of selection. The addresses within the boundaries of sample segments were then listed by trained interviewers, and served as the sampling frame from which the address sample for the NMES screener interview was selected. Approximately 35,000 addresses were selected for screening, within the sampled segments. Following the screening interview a subsample of dwelling units were selected for the full panel household survey from those screened, according to person and household level demographic characteristics. Subsampling rates were specified to obtain the required sample size to satisfy NMES precision specifications for person level estimates.

Screener Sample

As indicated, the 1987 NMES household survey was designed to provide an assessment of the annual health care utilization, expenditures, sources of payment and health insurance coverage in 1987, for the civilian non-institutionalized U.S. population. In addition, the NMES specifically targeted population subgroups of special policy interest for oversampling. More specifically, the NMES design required selective oversampling of blacks, Hispanics, the poor and near poor, those 65 years of age or older, and the functionally impaired.

The poverty status classification was based on the Poverty Index, developed at the Social Security Administration in 1964 and revised by the Federal Interagency Committees in 1969 and 1980. The poverty thresholds are updated every year to reflect changes in the Consumer Price Index. The poverty status classification is defined at the family level (or person level for unrelated individuals) and is a function of family size and, for single or two person families, age of household head. The near poor classification was defined to consist of families with family income above the poverty level, but less than 1.25 times the poverty level. The 1986 family incomes and poverty level thresholds were used for classification purposes. To facilitate the acquisition of relevant income data to allow for the poverty status classification, show cards were administered during the screener interview. A set of show cards were printed to reflect the different income thresholds for a poverty status classification of poor, near poor, or other income. The interviewer would then present the appropriate showcard to the household respondant, based on family size, to ascertain the applicable poverty status classification for the household.

The functionally impaired are defined as persons needing or receiving personal assistance or the Help of special equipment for a period of three months or more to perform one or more specifically defined activities of daily living (eating, transferring from bed or chair, dressing, bathing,) because of a health or physical problem.

A set of strict precision requirements were specified to allow for detailed analyses of these policy relevant population subgroups. These precision specifications are listed in Table 1, for a 20 percent person level estimate characterizing the respective analytical domains.

The 1977 NMCES consisted of 1,290 segments with an average segment size of 12.9 per segment. Design effects generally ranged from 2 to 5 for survey estimates (Cohen, 1983). The NMES requirement of an average design effect of 1.7 for survey estimates was specified to insure an efficient survey design. To satisfy this requirement, an average segment size of 6 households per segment was planned for NMES. Given the overall sample size requirement of 14,000 household interviews, 2,317 segments were selected, nearly twice the number considered in NMCES.

Due to the complex sampling requirements in the NMES, a separate screening interview was used to facilitate sample identification of population subgroups targeted for sampling at differential selection rates. The separate screening operation allowed for analyses of the results of the screener interview prior to the selection of the NMES Round 1 household sample. More specifically, it allowed for analysis of three features of the sample before the design was finalized: (1) a determination of whether the number of households screened was sufficient to provide the required sample sizes for each analytical domain; (2) whether the person based domain requirements were consistent with a sample size specification of 14,000 households completing the full series of interviews; and (3) development of the subsampling rates needed to attain the desired sample size in each domain. An advanced screening operation would make available precise counts of the outcome of the NMES subsampling process, prior to fielding the Round 1 household sample.

A self-weighting sample design was developed for the NMES screener interview to insure an efficient sample. The number of screening interviews required to meet targeted precision specifications was determined by identification of the demographic category which required the highest sampling rate for inclusion in the NMES household survey. The estimated sample size requirement for the NMES household screening interview was driven by the precision requirement for the black, 65 +, poor population subgroup. The households that completed the screening interview also served as a base from which individuals associated with the remaining demographic categories could be identified for inclusion in the NMES.

To illustrate the method of sample size determination for the NMES household screeners consider the black, 65 + poor subgroup, which was projected to require the largest number of completed screenings to yield the necessary number of respondents completing the full series of NMES interviews. To achieve a relative standard error of 18.0 percent on a 20 percent population estimate (the NMES requirement) for individuals who are black, 65 + and poor, the following sample size (n) is required, assuming a design effect of 1.7:

Table 1

HH6 Precision Specifications for Subdomains of Interest

Subdomain	March 1985 CPS estimate of Population size (in 1,000's)	Relative standard error on 20% estimate
Black	27,800	3.5
<65	25,575	3.5
65+	2,225	10.0
Poor	10,042	5.5
Near poor	1,960	12.0
Other	15,978	4.5
65+, poor	705	18.0
Hispanic	16,890	4.5
<65	16,072	4.5
65+	818	19.0
Poor	4,979	8.0
Near poor	1,273	15.0
Other	10,638	5.5
White	182,210	2.0
<65	158,920	2.0
65-79	18,876	4.5
80+	4,413	9.0
Poor	19,163	4.5
Near poor	7,845	7.0
Other	155,201	2.0
65+, poor	2,395	12.0
65+, near poor	1,896	14.0
Other race, poor	1,280	17.0
Other race, other	5,337	11.5
TOTAL	233,516	1.5
Functionally impaired, 65+	3,312	9.0
Functionally impaired, <65	1,379	14.0
Not Functionally impaired	228,825	
TOTAL	233,516	

$n = \frac{(100 - P)}{P(RSE)^2}$ [Design Effect], where P = 20 percent, Design Effect = 1.7, and the RSE is the Relative Standard Error = .18. Consequently, $n = \frac{80}{20(.18)^2} \times 1.7$, or n = 210. The required number of housing units needed to identify 210 black, 65+, poor respondents after Round 4 is derived by:

$$H = \frac{n}{2.7 \times \left\{\begin{array}{c}\text{proportion of U.S. civilian} \\ \text{noninstitutionalized population} \\ \text{who are black, 65+, and poor}\end{array}\right\} \times \left(\begin{array}{c}\text{Coverage} \\ \text{Rate}\end{array}\right)}$$

In this setting, n = 210, 2.7 is the number of respondents per housing unit, .00302 is the March 1985 Current Population Survey estimate of the proportion of the U.S. civilian noninstitutional population who are black, 65+ and poor, and .955 is the estimated Coverage Rate. Consequently:

$$H = \frac{210}{2.7 \times \left(\frac{705}{233,516}\right) \times (.955)} \qquad [= 26,976 = 27,000]$$

indicating that 27,000 dwelling units must complete the screener interview to satisfy NMES precision specifications.

Since the NMES precision specifications are based on individuals completing the full series of NMES household interviews (4 rounds of data collection), additional adjustments which control for survey nonresponse and vacancy rates had to be incorporated in the screener sample size specifications. Using the experience of Current Population Survey, the estimated occupancy rate in the U.S. was 92.5 percent. Further, the expected response rate for the screener interview was 95 percent, and the expected overall response rate for Rounds 1 through 4 of the household survey was 85 percent. Consequently, the number of addresses to be selected for the NMES screener sample was specified as 36,150, to obtain 210 respondents in the black, 65 + poor subgroup in round 4 (Table 2).

The targeted screener sample of 36,150 addresses was equally divided across data collection organizations requiring a selection of 18,075 sample addresses from the 1,150 sample segments that characterized the Westat national area probability sample, and a selection of 18,075 sample addresses from the 1,167 sample segments that characterized the NORC national area probability sample. For both organizations, the probability of selection is constant for all addresses sampled for the screener interview, regardless of the PSU or segment in which an address was located.

The screener sample consisted of dwelling units, although the basic analysis units are persons. The sample dwelling units (DU's) include housing units, group quarters, and other non-institutional (non-group) living quarters. All civilians who considered the selected DU as their usual place of residence were included in the interview.

A housing unit is defined as a house, apartment, group of rooms or a single room which is occupied as separate living quarters or vacant but intended for

Table 2

Sample size requirements, HHS

(1)	Number of addresses to be visited	36,150
	Expected occupancy rate	.925
(2)	Number of occupied dwelling units (eliminating vacant housing units and nondwelling units)	33,450
	Expected response rate for HHS screener	.95
(3)	Completed screening interviews required	31,800
	Expected response rate for HHS Rounds 1-4	.85
(4)	Number of households required after Round 4 to identify desired number of respondents in rarest population subgroup targeted for oversampling (this number determines the required number of screenings)	27,000

occupancy as a separate living quarters. The people who live there must eat and live separately from everyone else in the building, and the room or group of rooms must have either a separate entrance directly from the outside of the building or through a common hall, or complete kitchen facilities for the use of the household only. Complete kitchen facilities include all of the following: a range or cooking space (not a hot plate), a permanent sink with piped water, and a mechanical refrigerator.

A group quarter is defined as a single civilian non-institutional dwelling or structure in which five or more unrelated persons reside and where inhabitants are not considered a part of any other housing unit. Examples are boarding and rooming houses, hospital staff quarters, monasteries and convents, communes, non-military barracks, bunk houses, flop houses, and missions.

Included in a sample household within the dwelling units are persons who are considered to be a part of the household, but who are temporarily residing elsewhere. Unmarried college students under 22 were selected at their parents' sample addresses rather than at colleges in order to allow for the derivation of health care utilization and expenditure estimates at the family level. College students are interviewed at their college address when residing there.

Analysis is planned at both the individual and what is referred to as the reporting unit level. A reporting unit consists of one person living alone or unrelated to others in the same dwelling unit, and two or more persons related to each other by blood, marriage, adoption, or foster care whose usual residence is the assigned DU. More than one reporting unit may be living in the selected DU. Reporting units are comparable to census classifications of unrelated individuals plus families.

NORC selected its sample of 18,075 addresses after all the listing information was established on a computer data base. Due to time constraints, Westat selected its sample of addresses from listed segments on a flow basis as they became available. An overall sampling rate was chosen based on the estimated number of dwelling units for 1986. Each segment was then sampled individually at a rate which provided a constant overall sampling fraction. Since the total number of listed units was not known precisely in advance, the exact overall sample size could not be controlled.

Westat's initial sample selection consisted of 17,016 addresses. An additional 264 cases were added during Westat's quality control operations. The 4.5 percent short fall in sample size was primarily due to an over-estimate for 1986 of the number of dwelling units and a missed rate of 2.5 percent in the listing operation.

NMES Household Screening Interview - Field Results

The NMES screener interview was conducted during the fall of 1986 (October to December) at dwelling units selected from the respective Westat/NORC national area samples. An advanced letter describing the purpose of the NMES was mailed to each selected address. Interviewers visited each selected address to determine whether it was occupied as a primary residence, and attempted to conduct the interview with a household member. The household screener enumerated persons living in the selected dwelling unit, identified reporting units within the dwelling unit, and obtained demographic and other information to permit the subsampling of dwelling units to achieve the desired sample sizes for various domains of interest. The person identified as the reference person by the household member, and all other persons in the household related to the reference person formed a primary reporting unit (RU). Persons unrelated to the reference person also living in the sampled dwelling unit were grouped into other reporting units, with all related persons in the same reporting units. A separate screening interview was to be completed for each reporting unit, identified in a dwelling unit.

The final NMES household screener sample consisted of 35,634 addresses, of which 3,091 were identified as vacant and another 1,085 identified as not a dwelling unit. Of the 31,458 dwelling units eligible for the NMES, 28,708 responded to the screener interview. A dwelling unit was classified as responding if any of its component reporting units completed the screener interview. Consequently, the overall NMES household screener response rate was 91.3 percent. A more detailed breakdown of the final status classifications for the NMES screener sample is provided in Table 3.

NMES Round 1 Household Sample

The sample design had several features not usually required in area samples. First, initial precision requirements for the household survey were stated in terms of national estimates at the person level (Table 1). To meet or exceed these requirements, the survey must include at least the precalculated number of persons in each subdomain of interest. However, the unit of interviewing was the reporting unit and subsampling was done at the dwelling unit level. Thus, screened households were selected for the full panel HHS on the basis of the characteristics of the persons they included. Secondly, for both cost-efficiency, and to maximize the response rate, Round 1 was characterized by a

Table 3

NMES Household Screener Field Results

Total		Percent of Eligibles
Completed interview	28,708	91.60
Refusal	1,600	5.09
Screener Nonresponse (not at home after 4 calls, unavailable during field period, too ill)	916	2.91
Other Nonresponse (Language problem, unable to enter structure)	234	0.74
Eligible Total	31,458	100.00
Vacant	3,091	
Not a Dwelling Unit	1,085	
TOTAL	35,634	

"housing unit" design. That is, while households were selected based on the characteristics of their members, the Round 1 sample unit was the address of the household. The Round 1 interview was thus conducted with persons residing at the sample address at the time of the interview whether or not they lived there at the time of screening. These persons are then included throughout the full panel HHS, so long as they are in the civilian non-institutionalized population. Finally, a further complication was that subdomains are not mutually exclusive -- for example, a person can be both elderly and Hispanic. Consequently, in determining required sample sizes it was necessary to group the United States population into mutually exclusive cells by cross-classifying the variables that define the subdomains of interest. This cross-classification is presented in Table 4, in addition to the desired yields of sample respondents completing the full 4 rounds of NMES data collection, and their targeted subsampling rates.

The expected number of sample respondents at round 4 were derived by applying the approximate subsampling rates to a completed screener sample of 31800 households under the following assumptions: (1) 2.7 persons per household ,(2) an 85 percent overall Rounds 1-4 response rate, and (3) the March 1985 CPS estimated population distribution for the specified subdomains also characterized the 1987 distribution (Table 4).

To achieve a reduction in the overall costs of the NMES household survey, a modest reduction in the sample was implemented prior to the selection of the Round 1 sample. Generally, a six percent reduction in the Round 1 sample from

Table 4

NMES Sample Selection for HHS

Subdomain		March 1985 CPS estimate of population in subdomain (in 1,000's)	Proportion of total population	Approximate subsampling Rate	Expected Number of Round 4 Respondents
1.	Black, <65, poor	9,337	0.03998	8/9	2,591
2.	Black, <65, near poor	1,651	0.00707	8/9	458
3.	Black <65, other	14,587	0.06247	8/9	4,048
4.	Black, 65+, poor	705	0.00302	1	209
5.	Black, 65 near poor	309	0.00132	1	91
6.	Black, 65+, other	1,211	0.00519	1	359
7.	Hispanic, <65, poor	4,795	0.02054	8/9	131
8.	Hispanic, <65, near poor	1,168	0.00500	8/9	324
9.	Hispanic, <65, other	10,109	0.04329	8/9	2805
10.	Hispanic, 65+, poor	184	0.00079	8/9	51
11.	Hispanic, 65+, near poor	105	0.00045	8/9	29
12.	Hispanic, 65+, other	529	0.00226	8/9	147
13.	White, <65, poor	16,768	0.07181	6/9	3,490
14.	White, <65, near poor	5.950	0.02548	6/9	1,238
15.	White, <65, other	136,203	0.58327	3/9	14,173
16.	White, 65-79, poor	1,718	0.00736	6/9	358
17.	White, 80+, poor	677	0.00290	6/9	141
18.	White, 65-79, near poor	1,353	0.00579	6/9	282
19.	White, 80+, near poor	543	0.00233	6/9	113
20.	White, 65-79, other	15,805	0.06768	6/9	3,289
21.	White, 80+, other	3,193	0.01367	6/9	665
22.	Other races, poor	1,280	0.00548	6/9	266
23.	Other races, other income	5,337	0.02286	3/9	555
TOTAL		233,517	1.00000		37,013
					+ 352*
					37,365
24.	Functionally impaired, 65+	3,312	0.01418	8/9	919
25.	Functionally impaired, <65	1,379	0.00590	8/9	383
					1,302

*It is estimated that approximately 950 functionally impaired individuals would be obtained in the sample of 37,013 HHS Round 4 respondents, identified in subdomains 1-23, if no additional oversampling was considered for this subgroup. 352 functionally impaired individuals may need to be added. The subdomains will be further cross-classified on functional impairment status for sample selection.

initial specifications was implemented across each of the mutually exclusive population subdomains with one exception. An initial reduction of 16.67 percent in the Hispanic sample was specified, prior to the additional reduction of 6 percent. Consequently, the NMES Hispanic sample was reduced 21.67 percent from original specifications. In addition, the NMES sample of functionally impaired individuals 65 years of age or over was increased slightly over original specifications. Table 4 also presents the expected number of sample respondents at round 4, after implementing the sample reduction.

In the actual sample draw, households were selected sequentially. Households containing individuals targeted for selection at the highest rates were selected at the initial stages, one demographic category at a time. During the initial stages of sample selection, other members of sampled households who do not possess the demographic characteristic of the household classification were tallied in accordance with the remaining demographic categories, and the counts were subtracted from the sample quotas for the other domains. Thus, sampling rates for each category will change after each cycle of selection, from initial specifications based on the cumulative samples obtained for each category.

This sequential selection procedure permits control of sample sizes obtained for all demographic categories of interest so as to meet the specified precision levels. Because such a sequential procedure results in variation in sampling rates within demographic categories, there will be a slight increase in the variance of sample estimates attributable to the differential rates. However, for most groups this increase is quite small and has been accounted for in determining the sample sizes needed to meet the specified precision levels.

The NMES Round 1 household sample was selected after the data collection phase for the NMES screening interview was completed and all relevant demographic data necessary for the sample selection of responding dwelling units (race/ethnicity, age, functional impairment level, and income) was entered into the screener data base. To allow for the monitoring of sample yields from the NMES screener interview on an ongoing basis during the data collection phase of the screening operation, and to facilitate the availability of a complete screener data base in early January, 1987, the screener data was entered into a preliminary data base by a computer assisted data entry system upon receipt.

All eligible individuals within an eligible DU were assigned to one of the 23 mutually exclusive categories specified in Table 4 (and also to a functionally impaired category when appropriate). All individuals in the same reporting unit were given the same income classification. For the purposes of sample selection, an imputation strategy was considered for missing values at the person level for the following key sampling variables: age, functional impairment, poverty status, and income. The modal value for sample respondents was used as the imputed value for the NORC sample. The same strategy was considered for the Westat sample for age and functional impairment. Westat data at the segment level on race/ethnicity and income, used for sample selection, was also used to impute a value of race/ethnicity or poverty status for individuals with missing data who resided in the selected segment.

On a weekly basis between the months of October and December, 1986, monitor reports of screener interview sample yields were provided by Westat, which provided tabulations of the cumulative number of individuals falling into each sampling category, their distributions across all categories, and the total number of responding dwelling units.

Results of the NMES monitoring reports indicated a substantial departure from expectation in the poverty status distribution. In particular, sample yields for the near poor classes were generally larger than expected, and the corresponding yields for the other income classes were slightly smaller, while the "at or below poverty" classes showed an erratic pattern when compared to CPS estimates. The discrepancies in the income distribution appeared to reflect reporting errors rather than sampling problems. Since there were only small differences in the required sampling rates for many of the income classes within demographic groups, it was decided, for sampling purposes, to eliminate some of the domains. Income groups were therefore mostly collapsed within the following demographic groups: blacks, Hispanics, white 65-79, white 80+, and white less than 65 (poor/near poor). In addition, to minimize the effect of differential sampling rates on the precision of survey estimates, the "other races, near poor/other income" was combined with the "white, less than 65, other income" class. Prior to sample selection, dwelling units in the collapsed domains were sorted by income and race/ethnicity. Further, the separate domain for the functionally impaired, under 65 category was eliminated as a distinct sampling domain. As a result of collapsing the initial set of sampling domains, ten final sampling classes were specified.

The final ten collapsed sampling domains are presented in Table 5. With only one exception, dwelling units within each of these sampling classes were hierarchically sorted on the dimensions of race, poverty status, PSU, segment and household size. For the functionally impaired domain, dwelling units were only sorted by PSU, segment, and household size. A systematic sample of dwelling units was then selected within each of these classes. Since it seemed likely that some of the white persons under 65 who reported themselves as near poor actually had higher incomes, the targeted sample size for this domain was increased by approximately 300, to make sure the desired sample size was attained. The targeted sample for white, under 65, other income was reduced by the same margin.

The Round 1 household sample of dwelling units was then selected by the following sampling strategy (DiGaetano,1987):

1. Each screened DU was classified into one of the ten sampling categories based on the demographic characteristics of its "highest priority" individual. This was the household member requiring the highest sampling rate to meet sample size targets. The sampling rate was defined as the ratio of the required number of individuals in a sampling class to the available number of screener respondents. The sampling classes presented in Table 5 are arranged in order of highest priority.

2. A systematic sample of DU's was then selected from the highest priority class using the person level sampling rate.

3. The number of individuals in all sampled DU's that possessed the demographic characteristics that defined the remaining sampling classes were

Table 5 Final sample information: sample of completed screeners

Demographic group	Number of people available when completing sample for this group*	Total number of people selected for this group**	Number of DU's available for this group before any samples drawn	Number of DU's selected to complete sample	Sampling rate	Sampling interval
1. Blacks, under 65 years	7,735	7,735	2,730	2,730	1.0000	1.0000
2. Blacks, 65 years and over	467	731	361	325	.9014	1.1092
3. Functionally impaired, 65 years and over	1,110	1,148	1,054	943	.8945	1.1178
4. Hispanic, 65 years and over	217	206	174	136	.7788	1.2840
5. Hispanic, under 65 years	4,945	4,094	1,638	1,273	.7773	1.2864
6. Whites, 65-79 years	5,151	4,338	3,929	2,661	.6773	1.4763
7. White, 80 years and over	838	1,020	734	376	.5119	1.9533
8. White, under 65 years, or at below poverty and near poor	7,602	5,424	2,850	1,829	.6415	1.5587
9. Other races, at or below poverty	438	280	132	61	.4680	2.1365
10. White, under 65 years, other income Plus other races, near poor or other income	42,186	16,003	14,856	4,796	.3228	3.0975
TOTAL	–	39,831 √	28,458	15,130	–	–

* The numbers in this column represent the number of individuals in the associated demographic group when it came time to complete the sample for that group (other individuals in that group may have been in a DU associated with a higher priority group and thus were not available for sampling at the final opportunity for selection).

** The numbers in this column represent the total number of individuals selected who were specifically in that demographic group (they may have been selected with a DU at the final opportunity to select for that group or with a DU drawn for any higher priority group).

√ Among those people under 65 who were selected were 499 functionally impaired individuals between the ages of 18 and 64. No special sampling provisions were necessary for this group to obtain the desired sample size.

counted. The sample size requirements for the remaining sampling classes were then reduced by these enumerated totals.

4. DU's associated with the demographic category that was sampled were then removed from the sampling process.

5. DU's associated with the next highest priority sampling class were then selected at a sampling rate designed to meet sample size requirements, repeating the process outlined above.

6. The sampling process is repeated until all ten sampling categories have been sampled.

Sample of Addresses Vacant During the Screener Field Period

It was expected that some of the addresses contacted which were vacant in the screener field period would become occupied at the time of the Round One interview. Excluding them from the sample would understate the number of recent moves in the sample. Consequently, a sample of vacant addresses was selected to supplement the occupied addresses sampled from all screened households. Since eight percent of household addresses are expected to be vacant at any time, a sample of 1464 vacant addresses identified during the screening field period were selected, representing approximately eight percent of the total addresses to be visited in Round 1. It was assumed the number of vacant addresses becoming occupied would compensate for the number of occupied dwelling units becoming vacant in Round 1.

The set of 3,091 addresses identified as vacant during the screener field period were classified into one of four categories, based on the demographic characteristics of the sample segments from which the vacant addresses were drawn. The four classes were defined as black, Hispanic, white poor, and white other, reflecting the dominant characteristic of the individuals in the segments in which the vacant addresses were located. The sampling rates that were applied for the sample of vacant addresses were the minimum sampling rate associated with the stratum from the main sample draw. For example, if the category "Hispanic, under 65" had a lower sampling rate than "Hispanics, 65+", then the rate for 'Hispanics, under 65" was used to sample vacant addresses in segments classified in the Hispanic stratum.

Sample of Screener Refusals and Other Nonresponding Dwelling Units

A supplemental sample of refusals to the screener and other nonresponding dwelling units was also included in the sample, to improve the overall NMES response rate which reflects four rounds of data collection. In addition to 1600 dwelling units classified as screener interview refusals, another 916 dwelling units were categorized as "other nonresponse". The other nonresponse classification included dwelling units where no one was home after 4 calls, where the potential household respondents were unavailable during the screener field period, and where respondents were considered too ill to complete the interview.

Refusals and other nonresponding dwelling units were classified into the same four strata used for vacant addresses. A sample of 645 DU's categorized as refusals to the screener were then selected from the 1600 eligible DU's and a

sample of 376 DU's categorized as other nonresponding DU's were selected from 916 eligible DU's with this classification. Consequently, the only set of screener nonrespondents that were not considered for selection in the first round of the NMES household survey consisted of DU's with a language problem, DU's where the interviewer was unable to enter the structure, and the remaining nonresponse classifications. These 234 DU's were not targeted for selection in Round 1 because of the low probability of converting them to participate in the four rounds of the household survey. Theoretically, the potential response rate for the Round 1 household interview is 99.26 percent, if all screener refusals and other nonresponding DU's participated in this interview. The final Round 1 sample is summarized in Table 6.

Table 6

NMES Round One Sample

Final Status for Screener Interview	DU Totals	Percent of Eligible	NMES Round one Sample
Eligible DU's			
1. Responding to screener	28,458	91.19	15,130
2. Refusals	1,600	5.13	645
3. Nonrespondents (not home after 4 calls, unavailable during screener field period, too ill)	916	2.94	376
4. Remaining nonrespondents	234	.74	0
TOTAL	31,208	100.0	16,151
Vacant	3,091		1464

Characteristics of Screener Respondents

Demographic characteristics of the screener respondents were examined, to determine whether there were any systematic differences from the expected U.S. population distribution across the collapsed NMES sampling classes. An implicit assumption in deriving the expected U.S. population distribution for the specified sampling domains was that the population distribution reported in the March 1985 CPS was approximately the same in fall 1986. The estimated population distribution of respondents to the screener across sampling domains (excluding the functionally impaired category) and the distribution derived from CPS data are presented in Table 7.

Table 7

Demographic Characteristics of Screener Respondents

Demographic Group	Distribution of Screener Respondents	CPS Estimate of U.S. Population Distribution
Blacks, <65	.1015*	.1095
Blacks, 65+	.0102	.0095
Hispanics, 65+	.0034	.0035
Hispanics, <65	.0688	.0688
White, 65-79	.0802	.0806
White, 80+	.0198	.0189
White, <65, poor/near poor	.1108*	.0973
Other races, poor	.0075*	.0055
White <65 other income, Other races, near poor/ other income	.5978*	.6061
Total	1.0000	1.0000

Under the assumption of a design effect of 1.7 for the NMES screener sample, two-sided large sample Z tests were considered to determine whether significant differences existed in the demographic profiles of the screener respondants from the expected population distribution. In this setting the CPS estimates were treated as population parameters and an .05 level of significance was used for hypothesis testing.

The sample contained a lower proportion or blacks under the age of 65, than were in the Census Bureau's estimate of the U.S. population distribution. This finding is suggestive of either higher survey nonresponse or under-coverage of this population subgroup, which is consistent with the experience of most population surveys. Based on CPS estimates, the black, 65+, population subgroup was expected to require the highest sampling rate for inclusion in NMES, following the screener interview. The lower representation of blacks, under 65 as screener respondents required a re-ordering of the sample domains, shifting blacks under 65 to the highest priority classification.

In addition, the distribution of screener respondents was characterized by a higher representation of individuals classified as "white, under 65, poor and near poor", and "other races, poor", and a lower representation of "white under 65,other income/other races, near poor and other income", relative to

*Indicates a statistically significant difference at the .05 level.

the U.S. population distribution for these classifications. Rather than reflecting a higher nonresponse rate or under-coverage problem for "whites, under 65,other income", and for "other races, near poor/other income", this result may be a consequence of a misclassification error in the reporting of family income, which determined a household's poverty status classification. No other significant differentials were detected for the remaining demographic categories which characterized the NMES sampling domains.

Estimation and Analysis

The household survey component of the NMES was designed to yield approximately unbiased national and regional estimates of demographic and health care parameters for the 1987 noninstitutionalized population. Analyses are planned at the person level, at the family level, and at the health care event level (e.g. physician visits, hospital stays, use of prescribed medicines). The estimation strategy for the NMES household survey will include adjustments for all levels of nonresponse. Nonresponse adjustments will be made at the housing unit level, the reporting unit level, and the person level, for complete and partial nonresponse for a person's entire period of eligibility in 1987. To further improve the precision of survey estimates that characterize the NMES sample, a poststratification adjustment at the person level for age-race-sex cross-classifications will also be considered.

As noted, the NMES is characterized by a complex survey design which includes stratification, clustering, and disproportionate sampling. This departure from simple random sampling assumptions requires special consideration with regard to variance estimation and analysis. Several methods for approximating sampling variances, which incorporate the components of a complex survey design, have been developed and will be considered for application to NMES data (Cohen, Burt and Jones, 1986). The three most generally accepted and frequently used techniques are the method of balanced repeated replication (BRR), the "jackknife" method, and the Taylor series linearization method (Cohen and Kalsbeek, 1981). These variance estimation strategies have been incorporated as procedures in several of the widely used statistical packages. One can gain access to repeated replication procedures to compute estimates of regression coefficients and their sampling errors for data from complex survey designs through the use of the OSIRIS IV Statistical Analysis and Data Management Software Systems Package (Van Eck, 1979). The OSIRIS Repeated Replication Sampling Error Analysis procedure, REPERR, allows for the creation of replication by using one of three methods: balanced half-sample, jack-knife, or user-specified replications. In addition, the SAS accessible NASSREG procedure developed by Westat (Chu, et al., 1985) also uses the balanced half-sample replication technique for variance estimates of regression coefficients. The balanced half-sample method of variance estimation for means, proportions, totals, and rates can be implemented through the Health Examination Survey Variance and Cross-Tabulation Program developed by the National Center for Health Statistics (Jones, 1983). Appropriate variance estimates for means, proportions and totals using the Taylor series linearization method can be generated through the SESUDAAN program (Standard Errors Program for Computing of Standardized Rates From Sample Survey Data) developed by the Research Triangle Institute (Shah 1981a) and accessible through SAS. Similarly, variance estimates of ratios can be

generated through the SAS accessible procedure, RATIOEST (Standard Errors Program for Computing of Ratio Estimates From Sample Survey Data), also by using a linearization approximation (Shah, 1981b).

The OSIRIS IV procedure, PSALMS (Sampling Error Analysis), provides another mechanism for computing variance estimates for means, proportions, totals, and ratios derived from complex survey data by using the Taylor series linearization method. The same capabilities are shared by the SUPERCARP program (Cluster Analysis and Regression Program) developed by the Survey Section of the Statistical Laboratory at Iowa State University (Hidiroglou, Fuller, and Hickman, 1980). In addition, the procedure provides standard errors for regression coefficients, again by using a Taylor series linearization method. A complementary procedure, which also generates standard errors of regression coefficients by using complex survey data, is the SURREGR program (Standard Errors of Regression Coefficients From Sample Survey Data; Holt, 1977).

Summary

The complex survey design of the household component of the National Medical Expenditure Survey has been described in detail. Particular attention has been given to the disproportionate sampling strategy used to oversample blacks, Hispanics, the functionally impaired, the poor and near poor, and the elderly. Field results from the screener interview are also presented. In addition, the details of a sampling strategy are described, that augmented the NMES Round One sample with a representation of addresses vacant during the screener field period, of screener nonrespondents, and of screener refusals. A comparison of the demographic characteristics of screener respondents with Current Population Survey estimates for the nation is also provided. Further, appropriate methods of variance estimation which consider survey design complexities have been suggested for use.

References

Chu, A., L. Mohadger, D. Morganstern, and M. Rhoades (1985). NASSREG (National Accident Sampling System Regression). Internal Westat Report. Rockville, Maryland.

Cohen, S.B., and W.D. Kalsbeek (1981). National Medical Expenditure Survey: Estimation and Sampling Variances in the Household Survey. National Center for Health Services Research, Instruments and Procedures Series, No. 2. Department of Health and Human Services Publication No. (PHS) 81-3281. Washington, D.C.: Government Printing Office.

Cohen, S.B. (1983). Design Effect Variation in the National Medical Care Expenditure Survey. Proceedings of the Survey Research Section, 748-756.

Cox, B.G., and S.B. Cohen (1985). Methodological Issues for Health Care Surveys. New York, Basel: Marcel Dekker, Inc.

DiGaetano, R.G. (1987). Sampling Report for the Selection of Dwelling Units for the Household Survey of the NMES. NMES Report No. 1.00. Westat, Inc., Rockville, Maryland.

DiGaetano, R.G., and J. Waksberg (1986). Sample Design of the Household Survey (HHS) and the Survey of American Indians and Alaska Natives (SAIAN). NMES Report No. 1.46. Westat, Inc., Rockville, Maryland.

Hidiroglou, M.A., W.A. Fuller, and R. Hickman (1980). SUPERCARP (6th ed.). Ames, Iowa: Iowa State University, Survey Section, Statistical Laboratory.

Holt, M.M. (1977). "SURREGR: Standard Errors of Regression Coefficients From Sample Survey Data". Technical report, Research Triangle Park, North Carolina: Research Triangle Institute.

Jones, G.K. (1983). "HESBRR" (HES Variance and Crosstabulation Program, Version 3). Unpublished report, Hyattsville, Maryland: National Center for Health Statistics.

Kish, L. (1965) Survey Sampling. New York: John Wiley and Sons.

Shah, B.V. (1981a) "SESUDAAN: Standard Errors Program for Computing of Standardized Rates From Sample Survey Data". Technical report, Research Triangle Park, North Carolina: Research Triangle Institute.

Shah, B.V. (1981b). "RATIOEST: Standard Errors Program for Computing Ratio Estimates for Sample Survey Data". Technical report, Research Triangle Park, North Carolina: Research Triangle Institute.

**DEPARTMENT OF
HEALTH & HUMAN SERVICES**
Public Health Service
National Center for Health Services Research
and Health Care Technology Assessment
Parklawn Building, Room 18-12
Rockville MD 20857

Official Business
Penalty for Private Use $300

OASH